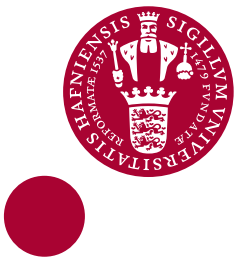

Circularity bias in abusive head trauma studies could be diminished with a new ranking scale

Högberg, Göran; Colville-Ebeling, Bonnie; Högberg, Ulf; Aspelin, Peter

Published in:
Egyptian Journal of Forensic Sciences

DOI:
10.1016/j.ejfs.2015.12.001

Publication date:
2016

Document version
Publisher's PDF, also known as Version of record

Egyptian Journal of Forensic Sciences (2016) 6, 6–10

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Egyptian Journal of Forensic Sciences

journal homepage: http://www.journals.elsevier.com/egyptian-journal-of-forensic-sciences

ORIGINAL ARTICLE

Circularity bias in abusive head trauma studies could be diminished with a new ranking scale

Göran Högberg [a,b,*], **Bonnie Colville-Ebeling** [c], **Ulf Högberg** [d], **Peter Aspelin** [e]

[a] *Department of Women's and Children's Health, Child and Adolescent Psychiatric Unit, Karolinska Institutet, Astrid Lindgren Children's Hospital, 171 76 Stockholm, Sweden*
[b] *Stockholm Child and Adolescent Psychiatry, Arenavägen 33[6], 12177 Stockholm, Sweden*
[c] *University of Copenhagen, Department of Forensic Medicine, Fredrik den V's Vej 11, 2100 Copenhagen, Denmark*
[d] *Department of Women's and Children's Health, Uppsala University, SE-751 85 Uppsala, Sweden*
[e] *Division of Medical Imaging and Technology, Department of Clinical Science, Intervention and Technology (CLINTEC), Karolinska Insitutet/Karolinska University Hospital, SE-14186 Stockholm, Sweden*

Received 7 September 2015; revised 22 November 2015; accepted 24 December 2015
Available online 14 January 2016

KEYWORDS

Abusive head trauma;
Circularity bias;
Ranking scale;
Caregiver

Abstract: Causality in abusive head trauma has never been fully established and hence no gold standard exists for the diagnosis. Implications hereof include bias introduced by circular reasoning and a shift from a trustful doctor patient relationship to a distrustful one when the caregiver statement is questioned. In this paper we examine seven recent abusive head trauma studies including 476 diagnosed abuse cases for circular reasoning as well as the role of the caregiver statement in the diagnosis. Secondly, we present a novel ranking scale for the diagnosis of abusive head trauma designed to minimize circular reasoning. We found circularity to be a potential source of bias in all seven studies. The caregiver statement (lack of trauma mechanism or trauma mechanism considered incompatible with clinical findings) was listed as a diagnostic item in 329 (69%) of 476 cases. Applying our ranking scale to the abuse cases showed that the demands of our ranking scale were not fulfilled in 440 (92%) cases. We conclude that most abuse cases in the studies were, to some extent, diagnosed on criteria based on circular reasoning. The caregiver statement was one of the most frequently used diagnostic items. Hypothetically, caregivers offer no or inadequate explanation to the clinical findings in assumed abuse cases. Thus, when this feature is encountered, it is regarded as indicative of abuse adding further to the risk of circularity bias.
We propose the use of our novel ranking scale in abusive head trauma research in an effort to minimize circular reasoning.

* Corresponding author at: Department of Women's and Children's Health, Child and Adolescent Psychiatric Unit, Karolinska Institutet, Astrid Lindgren Children's Hospital, 171 76 Stockholm, Sweden. Tel.: +46 736 27 19 19.
E-mail addresses: gorhogberg1@gmail.com (G. Högberg), bonnie.colville-ebeling@sund.ku.dk (B. Colville-Ebeling), ulf.hogberg@kbh.uu.se (U. Högberg), peter.aspelin@ki.se (P. Aspelin).

Peer review under responsibility of The International Association of Law and Forensic Sciences (IALFS).

http://dx.doi.org/10.1016/j.ejfs.2015.12.001

1. Introduction

In 1971 pediatric neurosurgeon Gutkelch hypothesized, that intracranial and intraocular bleeding in children with no external signs of injury was the result of shaking.[1] Radiologist Caffey subsequently published on the subject in a similar vein.[2] This hypothesis has been subjected to debate because the causal connection between exposure and clinical findings has never been established. It is unclear if shaking on its own is forceful enough to produce the clinical findings or if there is a lucid interval between exposure and onset of symptoms. Additionally, re-bleeding of a chronic subdural hematoma and strokes has been suggested as the cause of clinical findings in suspected abuse cases.[3–6] Despite these controversies it is widely accepted that the triad findings of subdural hemorrhage (SDH), retinal hemorrhage (RH) and encephalopathy are pathognomonic or highly specific for the syndrome known today as "Abusive Head Trauma" (AHT), so named by the AAP committee COCAN (American Academy of Pediatrics and Committee on Child Abuse and Neglect).[7]

Because the causal connection between exposure and clinical findings has never been established, no gold standard or standard case definition for AHT exists. Often clinical findings are occult and only investigated and disclosed when they are suspected and since the background population prevalence of these findings is unknown, the same applies to their predictive value. When there is no standard definition of abuse, it is often defined by the same variables that are subsequently analyzed as variables of abuse. Or, as stated by Piteau et al., "As there are no standardized criteria for the definition of abuse, most authors developed their own criteria, and many of these are fraught with circular reasoning".[8]

Recent attempts have been made to combine data from several observational studies to identify common diagnostic ground which has resulted in a substantial increase in clinical findings considered indicative of AHT.[8–10] These findings now include subdural hemorrhage, cerebral ischemia, retinal hemorrhages, skull fractures, intracranial injury, metaphyseal fractures, long bone fracture, rib fracture, seizure, apnea, bruising of the head, neck, ear and torso and no adequate history of trauma. The predictive value of this increasing list of findings is unknown and as long as research is fraught with circularity bias, the attempt to identify common diagnostic ground is more or less futile.

Circularity bias in AHT studies has received increased attention by several authors and different approaches have been suggested with the specific aim of avoiding circularity.[8,11,12] Among these is the ranking scale developed by Maguire et al.[9] (Box 1). However, "...for features that have been traditionally associated with abuse (such as subdural hemorrhage and retinal hemorrhage), this ranking scale does not compensate well for circularity" as suggested by Piteau et al.[8]

When the causative mechanisms in AHT are unclear, the evaluation of the proposed injury mechanism offered by the caregiver becomes important. Hypothetically, caregivers offer no or inadequate explanation to the clinical findings in assumed abuse cases. Thus, when this feature is encountered, it is regarded as indicative of abuse adding further to the risk of circular reasoning as well as creating an atmosphere of distrust in the doctor–patient relationship.

Box 1 Ranking scale for the diagnosis of AHT suggested by Maguire et al.[9]

Abuse ranking	Criteria used to define abuse
1	Abuse confirmed at case conference or civil, family, or criminal court proceedings or admitted by perpetrator or independently witnessed
2	Abuse confirmed by stated criteria, including multidisciplinary assessment
3	Abuse diagnosis defined by stated criteria
4	Abuse stated as occurring, but no supporting detail given as to how it was determined
5	Abuse stated simply as "suspected"; no details on whether it was confirmed

The objective of this short paper falls in two parts; firstly we examine recent AHT studies for possible bias caused by circular reasoning as well as the role of the caregiver statement in the abuse diagnosis. Secondly, we present a novel ranking scale designed specifically with the intent of avoiding circularity in AHT studies.

2. Methods and results

Two recent AHT studies were selected based on the study design and number of abuse cases. The studies' inclusion criteria were examined for circularity and checked against our own suggested ranking scale for abuse. The first study was conducted in 2011 by Maguire et al.[9] who selected 14 AHT studies identified as "high quality comparative studies". Of these, six entered the study based on data availability. Only cases of confirmed abuse as defined by the ranking scale previously presented by the authors were included. The other study was conducted in 2013 by Hymel et al.[10] who aimed at deriving a clinical prediction rule that, if validated, could be used as a tool for *excluding* AHT by identifying predictive clinical variables. They conducted their study from 14 pediatric intensive care units across the USA and the study population consisted of children less than three years old with acute head-injury admitted for intensive care. Children were categorized as abused by the following six definitional criteria which, according to the authors, were selected specifically to avoid circularity; admission by caregiver, independently witnessed abuse, caregiver denying head trauma, inconsistencies in the account by caregiver over time, caregiver account inconsistent with the developmental state of the child and presence of extracranial injuries considered suspicious of abuse. The following clinical variables were identified as predictive; acute respiratory compromise, seizures, bruising of ear, neck or torso, subdural hemorrhage and skull fracture.

The ranking scale we developed has three levels; ranking 1, 2 and 3. Ranking 1 is regarded as first grade evidence followed by ranking 2 and 3 in declining order.

Ranking 1: *Recorded*. With the introduction of smart cellphones follows a potential for recording everyday events in spontaneous home-videos. We consider it likely that the number of recorded cases of abuse (as well as accidents) of small children and infants will increase and we regard

this as first grade evidence. Additionally, it has the advantage of showing the biomechanics of the accident/injury.
Ranking 2: *Independently witnessed*. In witnessed abuse cases it must be clarified whether or not there is a conflict between the involved parties, e.g., a conflict over custody. As such, every witnessed case must be rigorously examined for bias caused by conflict of interest.
Ranking 3: *Confessed abuse*. It should be noted that this criterion is fraught with caveats. Firstly, it must be clarified what is being confessed to. Is it shaking of a baby in an effort to revive it or is it shaking of a baby in a fit of anger? How exactly was the baby shaken? There might be many different connotations to the concept of shaking. Secondly, are there any judicial consequences of the confession, like a plea bargain instead of a much longer imprisonment? Was the confession given at the presentation to the hospital or after police interrogation? The broad scope of problems with confessions in AHT cases has been presented by Deorah Tuerkheimer,[13] and the issue of producing false memories of committing a crime was presented by Shaw and Porter.[14] The often used Reid technique of interrogation based on psychological manipulation carries a clear risk of false confessions.[15]

Results are presented in Tables 1–3. Table 1 shows the proportion of cases with overlapping between diagnostic criteria and explanatory variables for AHT. There is a substantial overlap in regard to both caregiver statement as well as the classic triad findings. Table 2 shows the distribution of cases according to our ranking scale. There was a total of 476 abuse cases of which 34 were cases of confessed abuse (ranking 3) and two were cases of witnessed abuse (ranking 2). There were no cases of recorded abuse (ranking 1). Thus 92% of the 476 subjects diagnosed as abuse victims did not fulfill the demands of our ranking scale. Table 3 shows the proportion of cases in which the caregiver statement is a diagnostic item for abuse.

Table 2 Distribution of cases according to our ranking scale.

Study	Diagnosed abuse cases		Recorded	Witnessed	Confessed
	n	Age (years)	Rank 1	Rank2	Rank 3
Hettler and Greenes[16]	49	0–3	0	1	0
Kemp et al.[17]	65	0–2	0	0	19
Hobbs et al.[18]	106	0–2	0	0	0
Bechtel et al.[19]	15	0–2	0	0	1
Ettaro et al.[12]	89	0–3	0	0	0
Vinchon et al.[20]	57	0–2	0	0	0
Hymel et al.[10]	95	0–3	0	1	14
Sum all studies	476	0–3	0	2	34

The caregiver statement was listed as a diagnostic item in all seven studies and highlighted as contributory to the abuse diagnosis in 329 (69%) of the 476 cases. Five studies specified the caregiver statement as "no history, minor trauma or developmentally incompatible history".

3. Discussion

We found a substantial overlap between diagnostic criteria and explanatory variables in all seven studies and conclude that circular reasoning is a serious problem in AHT studies. There is a tendency to include infants with SDH in study populations and classify those with diagnostic criteria, primarily a rejected caregiver statement, as abused. Clinical findings in this group such as apnea, seizures, cerebral edema and other expected

Table 1 Diagnostic criteria for abusive head trauma. The proportion of cases with overlapping between diagnostic criteria and explanatory variables for individual variables.

Study	Inclusion criteria, *n*	Aim	Caregiver statement	Subdural hematoma	Retinal hemorrhage	Skeletal findings	Combination of extra-cranial findings
Bechtel et al.[19]	<2 yrs, *n* = 15, head injury with computed tomography	Identifying clinical variables	14/15	12/15	9/15		4/15
Ettaro et al.[12]	<3 yrs, *n* = 89, intracranial hemorrhage, skull fracture	Identifying clinical variables	86/89	58/89		28/89	
Hettler et al.[16]	<3 yrs, *n* = 49, intracranial hemorrhage	Diagnostic value of caregiver statement	47/49	43/49	39/49	6/49	
Hobbs et al.[18]	<2 yrs, *n* = 106, subdural hemorrhage/effusion	Identifying clinical variables	41/97	106/106	55/106	51/106	
Kemp et al.[17]	<2 yrs, *n* = 65, subdural hemorrhage/effusion	Apnea and brainswelling in AHT	21/65	65/65		31/65	
Vinchon et al.[20]	<2 yrs, *n* = 57, craniocerebral findings	Identifying clinical variables	57/57	46/57	42/57[a]		
Hymel et al.[10]	<3 yrs, *n* = 95, craniocerebral findings	Derive an AHT prediction rule from characteristics of AHT	63/95				53/95

[a] Authors comment: "In the construct of our study, however, we could not obviate the circularity bias, and the evaluation of the incidence of retinal hemorrhage in child abuse remains a self-fulfilling prophecy".

Table 3 The proportion of cases in which the caregiver statement is a diagnostic item for abuse.

Study	Diagnosed abuse cases		Caregiver statement diagnostic	Caregiver statement further specified as "No history/minor trauma or developmentally incompatible history"
	n	Age (years)	*n* (%)	
Hettler and Greenes[16]	49	0–3	47 (96%)	34/13
Kemp et al.[17]	65	0–2	21 (32%)	Not specified
Hobbs et al.[18]	106	0–2	41 (39%)	0/41
Bechtel et al.[19]	15	0–2	14 (93%)	12/2
Ettaro et al.[12]	89	0–3	86 (97%)	33/53
Vinchon et al.[20]	57	< 2	57 (100%)	Not specified
Maguire et al.[9]	381	0–3	266 (70%)	
Hymel et al.[10]	95	< 3	63 (66%)	32/31
Total sum	476		329 (69%)	111/140

findings in a brain injured infant are hereafter added to a list of diagnostic variables suspected of being associated with abuse. As commented by one of the authors of the studies selected for this study, "In the construct of our study, however, we could not obviate the circularity bias, and the evaluation of the incidence of retinal hemorrhage in child abuse remains a self-fulfilling prophecy".[11] The question is how many more clinical variables associated with abuse remain self-fulfilling prophecies in AHT?

Our suggested ranking scale was applied to seven AHT studies specifically aimed at avoiding circularity bias. According to our ranking scale, only 36 (8%) of 476 alleged abuse cases could be considered included in research reasonably free from circularity bias.

In 329 (69%) of 476 alleged abuse cases one of the diagnostic criteria was the rejection of the caregiver statement. Lack of trauma mechanism, minor trauma mechanism or trauma mechanism believed to be incompatible with the clinical findings is often interpreted as a sign of guilt. As stated by Hettler et al. "The most highly predictive historical feature for abuse is having no history of trauma"[16] and Vinchon et al.: "The medical diagnosis of child abuse was made in cases of traumatic lesions found in the absence of a history of a trauma or in cases in which the history of trauma changed from one report to another or was incompatible with the child's developmental age and/or the traumatic findings".[20] Thus, the decision by the physician to reject the narrative by the caregiver becomes a central element in the diagnostic process. This interpretation of the caregiver statement poses an inherent risk of misdiagnosis since all unwitnessed injuries will invariably be deemed suspicious of abuse when no trauma mechanism is offered. Furthermore, the criteria are flawed by circular reasoning. None of the examined studies have offered any reasoning for this assumption and we must conclude that there remains a need for scientific scrutiny of why the credibility of the caregivers is devaluated. The opposite hypothesis must also be tested; that the caregivers are truthful and that their narratives indicate a pathophysiological or accidental causation of the clinical findings. In this light, a minor trauma such as a short fall, which was reported by caregivers in 140 (56%) of 251 cases, suggests that such traumas might in fact produce the clinical findings. Additionally 44% of caregivers offered no trauma mechanism suggesting that an undiagnosed illness such as a venous thrombosis or rebleeding from a chronic subdural effusion caused the finding.

Circularity bias renders the diagnostic value of classical triad findings and diagnostic algorithms very low. This might explain the findings by Colville-Ebeling et al. who tested the diagnostic algorithms for abuse proposed by Hymel and Maguire on material from the National Inpatient Sample database (NIS KIDS).[21] This testing showed, that according to the Hymel algorithm every second brain injured child was to be considered abused. According to the Maguire algorithm, almost all brain injured children fit the definitional criteria of abuse. Such low predictability means that the rate of false positives is so high that the criterion-based abuse diagnosis is without meaning.

As the consequences of false positive as well as false negative abuse allegations are very harmful to the child, the whole family and society we consider it urgent that research in this field is carried out with a minimum of circularity bias. We thus propose that such research is sought to be conducted with cases selected according to our ranking scale for criteria to define abuse.

Funding

None.

Conflict of interest

None declared.

Ethical approval

Necessary ethical approval was obtained from the institute Ethics Committee

References

1. Gutkelch AN. Infantile subdural haematoma and its relationship to whiplash injuries. *Br Med J* 1971;**759**:430–1.
2. Caffey J. The whiplash shaken infant syndrome: manual shaking by the extremities with whiplash-induced intracranial and intraocular bleeding, linked with residual permanent brain damage and mental retardation. *Pediatrics* 1974;**54**:396–403.
3. Uscinski RH. Shaken baby syndrome: an odyssey. *Neurol Med Chir (Tokyo)* 2006;**46**:57–61.
4. Krous HF, Byard RW. Shaken Infant Syndrome: selected controversies. *Pediatr Dev Pathol* 1999;**2**:497–8.
5. Squier W. "Shaken baby syndrome" and forensic pathology. *Forensic Sci Med Pathol* 2014;**10**:248–50.
6. Geddes JF, Tasker RC, Hackshaw AK, Nickols CD, Adams HL, Whitwell HL, et al. Dural haemorrhage in non-traumatic infant deaths: does it explain the bleeding in "shaken baby syndrome"? *Neuropathol Appl Neurobiol* 2003;**29**(1):14–22.

7. Christian CW, Block RThe Committee on Child Abuse and Neglect. Abusive head trauma in infants. *Pediatrics* 2009;**123**:1409–11.
8. Piteau SJ, Ward MGJ, Barrowman NJ, Plint AM. Clinical and radiographic characteristics associated with abusive and nonabusive head trauma: a systematic review. *Pediatrics* 2012;**130**:315–23. http://dx.doi.org/10.1542/peds.2011-1545.
9. Maguire SA, Kemp AM, Lumb RC, Farewell DM. Estimating the probability of abusive head trauma. *Pediatrics* 2011;**128**(3): e550–64. http://dx.doi.org/10.1542/peds.2010-2949.
10. Hymel KP, Willson DF, Boos SC, Pullin DA, Homa K, Lorenz DJ, et al. Derivation of a clinical prediction rule for pediatric abusive head trauma. *Pediatr Crit Care Med* 2013;**14**:210–20.
11. Vinchon M, de Foort-Dhellemmes S, Desurmont M, Delestret I. Confessed abuse versus witnessed accidents in infants: comparison of clinical, radiological, and ophtalmological data in corroborated cases. *Child Nerv Syst* 2010;**26**:637–45. http://dx.doi.org/10.1007/s00381-009-1048.
12. Ettaro L, Berger RP, Songer T. Abusive head trauma in young children: characteristics and medical charges in a hospitalized population. *Child Abuse Negl* 2004;**28**:1099–111.
13. Tuerkheimer D. *Flawed convictions "Shaken baby syndrome" and the inertia of injustice*. New York: Oxford University Press; 2014 p. 97–126.
14. Shaw J, Porter S. Constructing rich false memories of committing crime. *Psychol Sci* 2015;1–11. http://dx.doi.org/10.1177/0956797614562862.
15. Kassin SM, Drizin SA, Grisso T, Gudjonsson GH, Leo RL, Redlich AD. Police-induced confessions: risk factors and recommendations. *Law Hum Behav* 2010;**34**:3–38.
16. Hettler J, Greenes DS. Can the initial history predict whether a child with a head injury history has been abused? *Pediatrics* 2003;**111**:602–7.
17. Kemp AM, Stoodley N, Cobley C, Coles L, Kemp KW. Apnoea and brain swelling in non-accidental head injury. *Arch Dis Child* 2003;**88**:472–6.
18. Hobbs C, Childs AM, Wynne J, Livingstone J, Seal A. Subdural haematoma and effusion in infancy: an epidemiological study. *Arch Dis Child* 2014;**90**:952–5.
19. Bechtel K, Stoessel K, Leventhal JM, Ogle E, Teague B, Lavietes S, et al. Characteristics that distinguish accidental from abusive injury in hospitalized young children with head trauma. *Pediatrics* 2004;**114**:165–8.
20. Vinchon M, Defoort-Dhellemmes S, Desurmont M, Dhellemmes P. Accidental and nonaccidental head injuries in infants: a prospective study. *J Neurosurg Pediatrics* 2005;**102**:380–4.
21. Colville-Ebeling C, Leith W, Freeman M. An evaluation of the discriminatory power of clinical findings of abusive head trauma using U.S. hospital data. In: *Proceedings of the 66th annual scientific meeting of the American Academy of Forensic Sciences*, Seattle, Washington, USA, February 17–22; 2014. p. 346–7.